基于Informer神经网络的农产品物流需求预测分析——以华中地区为例

左　敏，胡天宇，董　微*，张可心，张青川

（北京工商大学 电商与物流学院，北京 100048）

摘　要：保障农产品物流稳定性即是保障民生问题的关键。对农产品物流需求的预测是合理规划农产品物流稳定性的重要保证。然而，农产品物流需求的预测实际较为复杂，预测过程中会受到各种因素影响。因此，为了保证对农产品物流需求预测的准确性，需要考虑多方面影响因素。本研究以农产品物流需求作为研究对象，利用Informer神经网络构建预测农产品物流需求的神经网络模型，以华中地区河南省、湖北省和湖南省为例，对三省的农产品物流需求进行预测。同时用长短时记忆（Long Short-Term Memory，LSTM）网络和Transformer神经网络对华中三省农产品物流进行需求预测，将三种模型预测结果进行对比。对比结果表明本研究构建的基于Informer神经网络模型预测测试误差平均百分比为3.39%，低于LSTM和Transformer神经网络模型的4.43%和4.35%。并且用该Informer神经网络模型对三省预测出的预测值与实际值结果较为接近，河南省2021年的预测值为4185.33，实际值为4048.1，误差为3.389%；湖北省2021年的预测值为2503.64，实际值2421.78，误差为3.380%；湖南省2021年的预测值，2933.31，实际值为2836.86，误差为3.340%。表明该模型对华中三省的农产品物流需求预测的结果较为准确。三省2023年的预测值高于2021年的预测值。因此，在2021年物流运输配套设施的基础上，必须保证物流运输效率，加强物流运输能力，以满足华中地区日益增长的物流需求。

关键词：农产品物流；需求预测；Informer神经网络；长短时记忆神经网络；Transformer神经网络；深度学习

中图分类号：TP 393.01　　**文献标志码：**A　　**文章编号：**SA202302001

引用格式：左敏, 胡天宇, 董微, 张可心, 张青川 . 基于Informer神经网络的农产品物流需求预测分析——以华中地区为例[J]. 智慧农业(中英文), 2023, 5(1): 34-43.

ZUO Min, HU Tianyu, DONG Wei, ZHANG Kexin, ZHANG Qingchuan. Forecast and analysis of agricultural products logistics demand based on informer neural network: Take the central China aera as an example[J]. Smart Agriculture, 2023, 5(1): 34-43.

1　引　言

随着中国经济快速发展，人民的收入和消费能力不断提高，人们对农产品的需求也逐渐变得多样化，从原来吃得饱到现在吃得好，高品质的农产品越来越受到消费者的喜爱，这也就对农产品的物流服务质量提高了要求。

目前，农产品物流存在两个主要问题：首先，中国的农产品物流流量大，如何将不宜存放的农产品以最快的速度运输到消费者手上是当前物流的难题，尤其是在面临城市突发危机的情况下，不论是汽车、火车还是轮船和飞机等运输方式，运输量都有限，如何合理地规划农产品物流对于保障民生问题相当重要。其次，农产品物流具有很强的季节性和周期性，农产品成熟后，对物流具有很强的时效性要求。如瓜果蔬菜保质期较短，如果有限的物流运输能力不能得到合理的安排，将会导致农产品损耗，造成物质损失。因此，推动现代农产品流通体

收稿日期：2023-02-26

基金项目：国家重点研发计划项目（2021YFD2100605）

作者简介：左　敏，博士，教授，研究方向为食品安全大数据、农产品智能追溯、智能信息处理。E-mail：zuomin@btbu.edu.cn

*通信作者：董　微，博士，讲师，研究方向为食品安全大数据、自然语言处理、深度学习。E-mail：dongwei2019@btbu.edu.cn

系建设，是促进农产品保值增值、农业提质增效、农民增收致富的重要保障。准确预测农产品物流运输需求有利于制定合理的农产品物流规划，确保经济的长期可持续增长，提高人民的生活水平。

近年来，国内外诸多学者在物流需求预测上提出了富有创新性和实践意义的方法。姚冠新等[1]在2021年选取兴化市2004—2018年数据，利用灰色预测模型GM（1，1）对兴化市2019—2025年农产品物流需求进行了预测，结果表明兴化市农产品物流需求量整体呈上升趋势，2025年农产品物流需求量预计287.32万吨。张继良[2]利用灰色关联法对经济发展水平、产业结构、物流发展、人文因素、市场因素和区位优势指标进行分析并构建了物流需求体系，之后利用山东省2008—2019年生鲜农产品商品化数量及相关因素构建了多变量灰色预测模型、优化后的神经网络模型和主成分回归模型三种模型并结合Shapley值组合法对山东2020—2025年生鲜农产品物流需求进行了预测，选取的组合预测模型所得出各个年份的预测误差相对较为平缓，预测结果更优，准确性更高。组合模型2020—2025年预测值分别为14575.61、14626.89、14677.75、14782.28和14778.97。王晓平和闫飞[3]构造了遗传BP神经网络模型（Genetic Algorithm-Back Propagation，GA-BP），利用遗传算法[4]的全局搜索能力，以北京为例对城镇农产品冷链物流需求进行了预测，结果表明该模型对冷链物流需求与实际影响因素的非线性关系有较强的刻画和预测能力。皇甫红姣[5]通过对绵阳市生鲜农产品冷链物流展开实地调研，进而确定出与冷链物流密切相关的6种农产品，即水果、蔬菜、水产品、肉类、奶类、禽蛋，采用灰色预测模型对绵阳市2020—2029年的农产品产量进行了预测，2020—2029年水果预测值为26.92、29.23、31.76、34.51、37.53、40.81、44.42、48.34、52.63和57.32；蔬菜预测值为47.73、47.36、47.1、46.93、46.88、46.92、47.1、47.41、47.87和48.49；奶类预测值为11.27、13.2、15.67、18.82、22.77、27.73、33.94、41.73、51.48和63.73；肉禽蛋水产品预测值为36.42、37.99、39.97、42.41、45.41、49.06、53.53、58.99、65.65和73.85。刘文博[6]基于辽宁省生鲜农产品冷链物流需求预测研究的发展现状，选取2012—2019年生鲜农产品产量并采用灰色预测模型GM（1，1）对辽宁省生鲜农产品产量数据进行预测研究，其预测结果精度较高，能够反映出2020—2024年辽宁省的生鲜农产品需求量变化趋势。Wang和Li[7]基于GM（1，1）模型对我国农产品冷链物流需求进行预测，选取2014—2019年中国农产品冷链物流数据，对未来五年中国农产品冷链物流需求进行预测，预测2020—2025年需求量分别为278,126,650、34,211,034、420,813,365、5,176,283和63,670,320吨。该预测的实验结果可以为相关部门提供数据参考并促进农产品冷链供需平衡提供理论支持。Huang等[8]运用GM（1，1）模型和BP神经网络模型对广东省2000—2019年的物流需求进行了模拟和预测。结果表明，广东省物流需求预测指标体系具有良好的适用性。对比两种模型各20年预测结果得出结论，与GM（1，1）模型相比BP神经网络模型预测误差更小，预测结果更稳定。基于研究结果，作者认为电子商务平台和物流企业应重视对区域物流需求的预测，选择科学的预测方法，并鼓励实施新的配送模式。徐晓燕等[9]选择了小波神经网络、人工神经网络、遗传算法优化神经网络、粒子群优化神经网络（Particle Swarm Optimization-Back Propagation，PSO-BP）和长短时记忆网络（Long Short-Term Memory，LSTM）等5种方法来对山东省物流需求进行预测，得到的结果与真实值相比，LSTM预测精度更高，效果最好。Van der Laan等[10]基于无国界医生组织阿姆斯特丹运营中心（MSF-OCA）关于19个长期援助项目和2000多种医疗用品在2013年消耗的标准化消耗数据，对订单计划流程进行了分析预测，得出了人道后勤组织进一步改进的方向。Lu和Park[11]结合空间经济学、产业经济学和新古典经济学的理论，以鄂尔多斯市为研究对象，确定了能够生产大量货物的主导产业，并进一步预测了鄂尔多斯及其腹地的静态物流生成。该研究通过整合影响区域物流需求的各种因素，从空间经济原理的角度建立了物流需求潜力模型，并将物流需求预测的方法从单一的统计原理扩展到特殊和区域经济学的一个新领域。Baisariyev等[12]为了实现正确预测持续供应给意外维护操

作的备件的必要数量。选取Bootstrap方法并使用航空物流的实际数据对其性能进行了评估。分析表明，Bootstrap方法在航空备件预测中比其他方法更精确，可以充分满足所有类别的需求。Zeng等[13]利用灰色预测模型对广东省农村物流需求进行预测。结果表明，该预测模型具有良好的匹配性和精度。研究还发现，在中短期内，广东省对农村物流的需求将普遍上升。根据本研究的发现，政府应刺激农村消费品需求，鼓励技术创新，发展互联网-农村物流模式，并根据各地区的情况满足农村物流的供应要求。

对农产品物流需求的预测实际上会受到多种因素的影响，这也对预测增大了难度。现有研究针对多种影响因素主要是采用回归分析，然而由于影响因素的数据量较大且基于的历史数据较多，预测结果不理想。目前诸多研究均采用了长时间序列的历史数据作为预测基础，对时间序列进分析。时间序列分析常用于揭示某一现象的发展变化规律或从动态的角度刻画某一现象与其他现象之间的内在数量关系及其变化的规律性，被广泛用于水文预报[14-16]、环境污染控制[17, 18]、天文学[19, 20]和海洋学[21, 22]等方面的研究，而这一特性同时也能够满足农产品物流需求的预测。

华中地区作为中国中部物流枢纽，需要发挥承东启西、贯通南北的作用。当前华中地区为满足省间物流需求，付出的物流力量较多、成本较高以及承受压较大等问题。针对上述问题，本研究以华中地区河南省、湖北省和湖南省为例研究农产品物流需求，选取了较为全面的影响因素，采用了基于神经网络的时间序列分析的方式对未来农产品物流需求进行精准预测，从而减少各省物流力量的浪费。

2 研究数据

2.1 数据来源

选取2017—2021年《河南统计年鉴》《湖南统计年鉴》《湖北统计年鉴》以及2021年上述三省国民经济和社会发展统计公报作为数据来源。从数据源中收集并筛选统计数据，为预测三省农产品物流需求提供数据支撑。

2.2 数据提取

农产品的物流需求预测需要综合考虑多种影响因素。在分析2017—2021年实际数据的基础上，筛选出影响三省农产品物流需求的主要影响因素。主要影响因素的选取需要考虑当地的经济发展水平、人均消费能力、产业结构水平、物流发展水平、农产品供给因素和人文因素等。本研究选取影响因素包括地区生产总值G1（亿元）、居民人均可支配收入G2（元）、居民人均消费支出G3（元）、第一产业增加值G4（亿元）、第二产业增加值G5（亿元）、第三产业增加值G6（亿元）、农作物总播种面积G7（千公顷）、粮食产量G8（万吨公里）、人均粮食产量G9（公斤）、农产品年产量G10（万吨）、种植和养殖面积G11（千公顷）、货物运输量G12（万吨）、铁路营业里程G13（万公里），以及公路里程G14（万公里）。

选取粮食、肉、蛋、奶、油料、茶叶、水果、林产品、水产品等9类农产品。其中，林产品由苗木和原木组成，农产品年产量由9类农产品的年产量相加得到，木材产量单位为万立方，转化系数为0.8，即每万立方木材等于0.8万吨[23]。种植和养殖面积由农作物播种总面积、果园面积、本年采摘茶园面积、水产品养殖面积和油料播种面积相加得到。

以各省常住人口数乘以各省人均农产品消费量作为因变量S，即农产品的物流需求[24]。农产品的人均消费量是9种农产品的人均消费量之和，计算方法公式（1）所示。

$$S = X \times Y \tag{1}$$

其中，X为各省的年末常住人口，万人；Y为各省的人均农产品消费量，kg。

各省的人均农产品消费量由人均粮食消费量、人均食用油消费量、人均蔬菜及食用菌消费量、人均肉类消费量、人均禽类消费量、人均水产品消费量、人均蛋类消费量、人均奶类消费量、人均干鲜瓜果类消费量和人均食糖消费量十项相加组成。

以农产品物流需求高度相关的影响因素为自变量，可以分为经济发展水平、产业结构水平、物流发展水平、农产品供给因素和人文因素五大类。经济发展水平包括地区生产总值G1（亿元）和居民

人均可支配收入G2（元）；产业结构水平包括第一产业增加值G4（亿元）、第二产业增加值G5（亿元）和第三产业增加值G6（亿元）；物流发展水平包括货物运输量G12（万吨）、铁路营业里程G13（万公里）和公路里程G14（万公里）；农产品供给因素包括农作物总播种面积G7（千公顷）、粮食产量G8（万吨公里）、人均粮食产量G9（公斤）、农产品年产量G10（万吨）以及种植和养殖面积G11（千公顷）；人文因素包括居民人均消费支出G3（元）、各省的年末常住人口X（万）和各省的人均农产品消费量Y（千克），共计16个指标。进而建立华中三省农产品物流需求的预测模型。数据集样本如表1所示。

表1　影响华中三省农产品物流需求的指标样本数据集

Table 1 Sample data sets of indicators affecting the logistics demand of agricultural products in three provinces of Central China

年份	省份	G1/亿元	G2/元	G3/元	G4/亿元	G5/亿元	G6/亿元	G7/千公顷	G8/万吨公里	G9/公斤	G10/万吨	G11/千公顷	G12/万吨	G13/万公里	G14/万公里	S/万吨	X/万	Y/千克
	河南	44824.9	20170	13730	4139.3	20940.3	19745.3	14732.53	6524.25	683.49	11271.39	16670.25	230114	0.54	26.78	3215.1	9829	327.1
2017	湖南	33828.1	23103	17160	2998.4	13459.8	17369.9	8321.97	3073.6	449.29	5432.21	10248.41	225551	0.47	23.97	2562.3	6633	386.3
	湖北	37235	23757	16938	3529	15713.9	17992.2	7956.14	2846.13	482.93	5374.29	9796.19	188107	0.42	26.95	2030.4	5904	343.9
	河南	49935.9	21964	15169	4311.1	22038.6	23586.2	14783.35	6648.91	693.9	11369.49	16776.68	259884	0.54	26.86	3231.4	9864	327.6
2018	湖南	36329.7	25241	18808	3084.2	13904.1	19341.4	8111.09	3022.90	439.41	5424.7	10095.08	229957	0.51	24.01	2483.5	6635	374.3
	湖北	42022	25815	19538	3548.2	17573.9	20899.9	7952.9	2839.47	480.49	5414.61	9807.57	204307	0.43	27.50	2015.3	5917	340.6
	河南	53717.8	23903	16332	4635.7	23035.6	26046.5	14713.98	6695.36	695.81	11447.8	16914.99	219024	0.65	26.98	3464.4	9901	349.9
2019	湖南	39894.1	27680	20479	3647.2	15401.7	20845.2	8122.79	2974.84	430.6	4730.02	10581.25	189740	0.56	24.06	2580.97	6640	388.7
	湖北	45429	28319	21567	3809.4	18723	22896.5	7815.89	2724.98	460.15	5338.46	10271.75	188133	0.52	28.9	2034.74	5927	343.3
	河南	54259.4	24810	16143	5354	22220.9	26684.5	14687.99	6825.8	688.02	11584.08	16737.73	219939	0.65	27.03	3777.58	9941	380
2020	湖南	41542.6	29380	20998	4240.7	15949.2	21352.7	8400.13	3015.12	453.91	4440.67	10409.77	200878	0.56	24.11	2700.53	6645	406.4
	湖北	43004.5	27881	19246	4133.2	15933.8	22937.6	7974.39	2727.43	467.35	5339.24	9753.15	160422	0.52	282.96	2205.51	5745	383.9
	河南	58887.4	26811	18391	5620.8	24331.6	28934.9	14705.13	6544.17	660.23	11282.39	16713.72	255551	0.65	27.16	4048.1	9883	409.6
2021	湖南	46063.1	31993	22798	4322.9	18126.1	23614.1	8504.26	3074.36	463.46	4701.37	10552.7	224465	0.56	24.19	2836.86	6622	428.4
	湖北	50012.9	30829	23846	4661.7	18952.9	26398.4	8109.24	2764.33	477.64	5589.54	9949.02	214762	0.52	29.69	2421.78	5830	415.4

2.3 数据规范化

由于每类影响因素的输入数据值在数量级上相差较大，不同的输入数据具有不同的量纲，为了消除农产品物流需求相关因素之间的量纲影响，需要进行数据标准化处理，以解决数据指标之间的可比性。因此，本研究对农产品物流需求相关因素采用最大-最小规范化处理，消除奇异样本数据导致的不良影响，归一化[25]操作如公式（2）所示。

$$G' = (G - G_{min}) / (G_{max} - G_{min}) \quad (2)$$

其中，G'是归一化的数据；G是原始数据；G_{max}和G_{min}分别是原始数据中的最大值和最小值。

通过对2017—2021年的华中三省农产品物流需求数据进行清洗等预处理后，计算各省份每年农产品物流需求总量S，并对数据进行归一化处理，从而完成农产品物流需求数据集的构建。

3　预测模型构建

深度学习算法已被广泛应用于农业病害识别、医疗诊断、地质勘探、工业设备故障诊断中。然而在时间序列预测问题中，序列的分布可能随时间轴的推进不断变化，这就需要模型具备更强的外推能力。Informer是Transformer改进的神经网络模型，通过使用多头概率稀疏自注意力机制，给重要性特征分配更大的权重，降低了网络模型的时间复杂度，同时基于生成式解码提高预测速度，解决了时间序列数据的长时依赖问题，有着较好的外推能力。因此，本研究选用Informer神经网络模型对华中三省农产品物流需求进行预测。

图1是华中三省农产品物流需求预测的Inform-

er神经网络模型，该模型由编码器和解码器两个部分组成。Informer模型通过多头概率稀疏自注意力层、多头注意力机制、掩盖多头概率稀疏自注意力层完成编码和解码器的构造，注意力机制从本质上讲和人类的选择性视觉注意力机制类似，目的也是从众多信息中选择出对当前任务目标更关键的信息，从而解决长期时序预测问题，提高预测模型的性能。

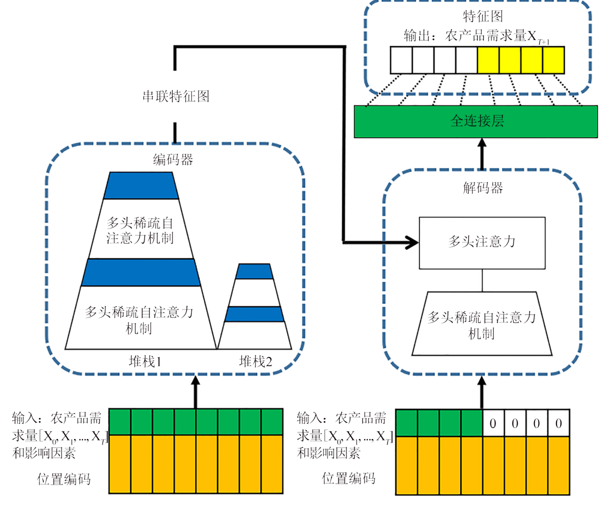


图1 基于Informer神经网络的华中地区农产品物流需求预测模型

Fig.1 Forecasting model of agricultural products logistics demand in three provinces of Central China based on informer neural network

在构建好的模型里输入历史农产品需求量和已选取的影响农产品物流需求的指标作为Informer网络的输入，如农产品年产量、公路运输、铁路运输等指标。使用位置编码标记时间序列$[X_t, \cdots, X_{t+i}, \cdots, X_T]$局部和全局前后时间位置关系，时序此时具有一定位置编码作用，但是位置编码标记的时序存在可以加快充分挖掘时间序列的时间相关性的速度，该位置编码是Prophet经典传统算法。利用Informer的多头注意力机制将注意力聚焦于指标的数据特征上，从而获取影响农产品物流需求的长时依赖关系。解码器输入由两部分组成，其中一部分为编码器，输出关于华中三省农产品物流需求风险的特征数据；另一部分将待预测的华中三省农产品物流需求量在输入时使用0进行占位，并添加掩盖机制。将数据连接到多头注意力机制，进而连接一个全连接层输出华中三省农产品物流需求预测数据X_{T+1}，从而实现对华中三省农产品物流需求的预测。

注意力机制计算如公式（3）所示，其中包含3个向量，分别为Query向量（$\boldsymbol{Q} \in R^{L_Q \times d}$）、Key向量（$\boldsymbol{K} \in R^{L_K \times d}$）和Value向量（$\boldsymbol{V} \in R^{L_V \times d}$），其中，$L_Q$、$L_K$和$L_V$分别表示Informer模块中查询集、键集和值集的线性变换层，d为输入维数。设q_i、k_i、v_i分别是$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$的第i行，则第i个Query的注意力被定义为概率形式的核平滑器。

$$A\left(q_i, \boldsymbol{K}, \boldsymbol{V}\right) = \sum_j \frac{k(q_i, k_j)}{\sum_l k\left(q_i, k_l\right)} v_j = E_{p\left(k_j | q_i\right)\left[V_j\right]} \tag{3}$$

第i个Query的稀疏度量为公式（4）。

$$M(q_i, \boldsymbol{K}) = \ln \sum_{j=1}^{L_k} e^{\frac{q_i k_j^{\mathrm{T}}}{\sqrt{d}}} - \frac{1}{L_k} \sum_{j=1}^{L_k} \frac{q_i k_j^{\mathrm{T}}}{\sqrt{d}} \tag{4}$$

基于提出的度量，通过允许每个键只关注u个主要查询来获得ProbSparse自我注意力，如公式（5）所示，其中$\bar{\boldsymbol{Q}}$为前u个Query的稀疏矩阵。

$$\mathrm{Attention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \mathrm{softmax}\left(\frac{\bar{\boldsymbol{Q}} K^t}{\sqrt{d}}\right) \boldsymbol{V} \tag{5}$$

Informer编码器负责将输入的华中三省农产品物流需求指标时间序列进行编码，获取三省农产品物流需求指标和影响因素时间序列的依赖关系并映射为需求指标信息的中间特征，其内部结构由两个结构相同的stack组成，stack结构如图2所示。Stack1是主堆栈，接收整个输入序列，而Stack2与Stack1操作相同，但是只接收一半的输入切片。每个堆栈均由编码层和蒸馏层组成。其中，编码层包括了多头概率稀疏自注意力层、前向神经网络、残差连接和正则化运算，如公式（6）所示。Informer内stack结构中的蒸馏层则是利用蒸馏机制来提升网络的鲁棒性，这是informer内stack结构的一个特征，因此能够降低网络的使用内存。最后，将所有堆栈的输出串联起来，得到encoder的最终隐藏表示，即图1中所示的特征图。

$$O = \mathrm{LayerNorm}(x + \mathrm{Sublayer}(x)) \tag{6}$$

其中，O表示对每个子层采用了层归一化操作。Sublayer为多头稀疏自注意力机制和前向神经网络处理函数；LayerNorm为正则化函数。

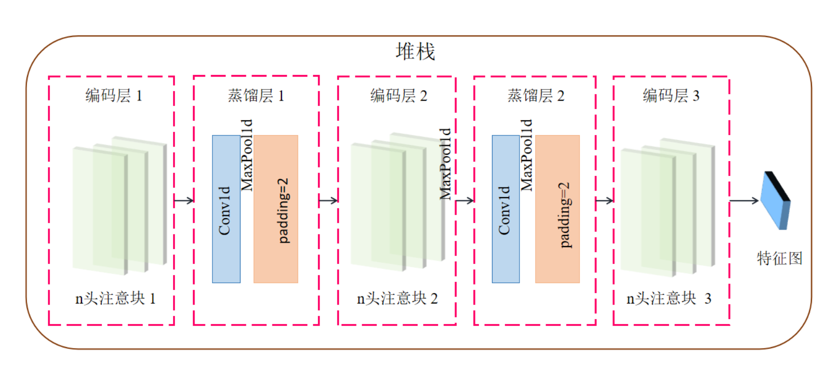


图 2 Stack 结构图

Fig.2 Structure diagram of stack

蒸馏机制即在时间维度上使用一维卷积操作，并通过 ELU 激活函数后添加池化层实现输入长度减半，蒸馏层比编码层少一层。蒸馏操作从第 j 层推进到 $j+1$ 层的过程为公式（7）：

$$x_{j+1} = \mathrm{MaxPool}(\mathrm{ELU}(\mathrm{Conv1d}([x_j]_{AB}))) \quad (7)$$

其中，[·]表示包括多头概率稀疏自注意操作和注意力块的关键操作；Conv1d 表示一维卷积操作；MaxPool 为最大池化操作；ELU 为激活函数，其计算方法为公式（8）。

$$\mathrm{ELU}(x) = \begin{cases} e^x - 1, & x < 0 \\ x, & x \geqslant 0 \end{cases} \quad (8)$$

4 实验与结果分析

本研究利用构建的 Informer 模型以及选取的 LSTM 和 Transformer 模型对华中地区农产品物流需求进行预测，进而对比三种模型的预测效果。选取预测效果较好的模型对华中地区 2023 年物流需求进行预测，得到的预测值作为参考，为河南省，湖南省和湖北省农产品物流的规划提供依据。

4.1 实验方法

采用 PyTorch 开源的深度学习框架作为实验平台，具体实验工具包和参数如表 2 所示。Informer 神经网络用于预测华中三省的农产品物流需求，使用 5 折交叉验证法进行模型训练。具体地，原始数据集被分为 5 个近似的平均集，在每个训练过程中，选择 4 个集作为训练集，其余 1 个集作为测试集。训练重复 5 次，每次选择的训练集和测试集都不同。

表 2 华中地区农产品物流需求预测实验环境参数

Table 2 Experimental environmental parameters of agricultural products logistics demand forecast in Central China

环境类型	环境名称	参数
操作系统	Windows 10	64 bit
硬件信息	CPU	Intel(R) Core(TM) i5-8265U CPU @ 1.60 GHz (8 CPUs)～1.8 GHz
	GPU	Radeon 540X Series
	RAM	16 GB
软件工具	Python 3.7	Numpy 1.21.6
		Scikit_Learn 1.0.2
		Pandas 1.3.5
		Torch 1.12.1
		Matplotlib 3.5.3

4.2 结果分析

本研究构建的 Informer 神经网络如图 1 所示。历史农产品需求量和已选取影响农产品物流需求的 16 个影响因素指标和省份、年份作为网络的输入，需要预测的对应省份的农产品需求量（万吨）作为输出。选取 LSTM 和 Tranformer 作为 Informer 模型的对照组。LSTM 是 RNN 的一个优秀的变种模型，继承了大部分 RNN 模型的特性，同时解决了梯度反传过程由于逐步缩减而产生的消失梯度问题，能够很好地预测长时序问题。Transformer 可以直接计算每个词之间的相关性，不需要通过隐藏层传递并且可以并行计算，可以充分利用 GPU 资源。故选取二者作为实验对照。Informer、LSTM 和 Transformer 训练结果分别如图 3～5 所示。

对比图 3～5 可以看出，Informer 模型的预测效果优于 LSTM 和 Transformer 模型，Informer 模型选

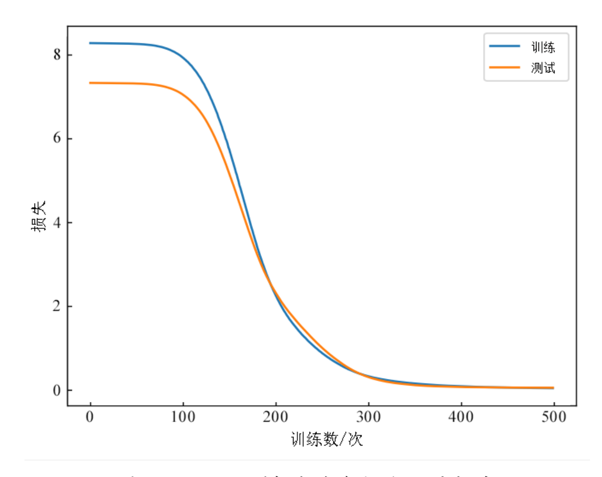


图 3　Informer模型训练和验证的损失

Fig. 3　Loss of training and verification of Informer model

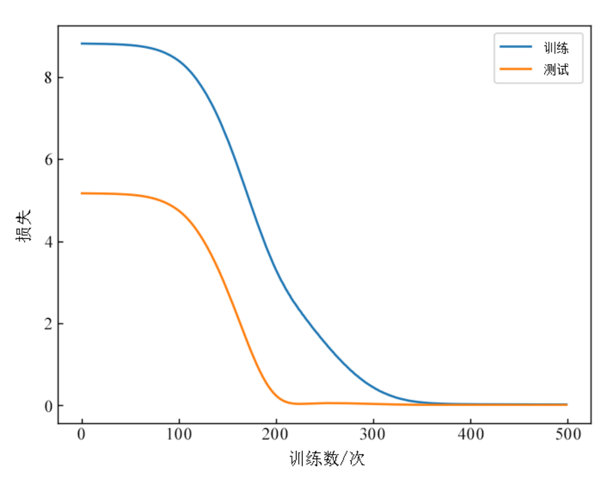


图 4　LSTM模型训练和验证的损失

Fig. 4　Loss of training and verification of LSTM model

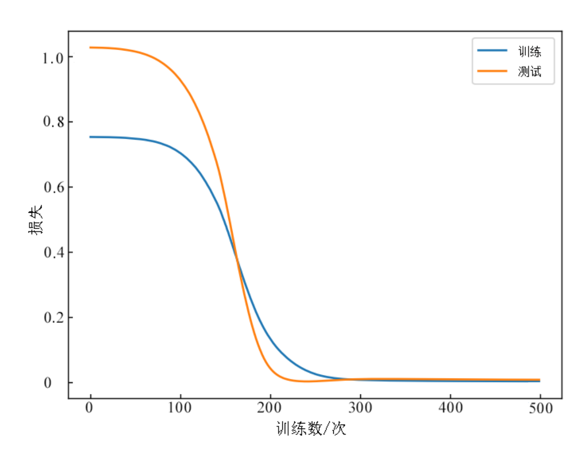


图 5　Transformer模型训练和验证的损失

Fig.5　Loss of training and verification of Transformer model

代刚开始时损失函数值略大，在迭代过程中会逐渐下降，然后逐渐趋于平缓。当迭代次数为180次时，损失函数值接近于0。LSTM模型预测效果较差，迭代初期损失函数值较大，当迭代次数为350次时，损失函数值才接近于0。Transformer模型预测效果一般，迭代次数为150次左右时，损失函数值接近于0，之后损失函数值增加，到迭代次数280次左右时再次损失函数值接近于0并逐渐趋于平缓。

表3显示了每个训练集和测试集的绝对误差和百分比误差，以及五重交叉验证的平均绝对误差和平均百分比误差。

表 3　神经网络的训练和测试误差

Table 3　Training and test errors of neural network

折叠数量	训练错误		测试错误	
	绝对误差	误差百分比/%	绝对误差	误差百分比/%
Informer 1倍	77.21	4.61	85.96	5.17
Informer 2倍	21.49	2.17	31.19	2.34
Informer 3倍	20.87	2.23	33.47	2.53
Informer 4倍	32.94	3.54	37.26	3.59
Informer 5倍	30.37	3.43	36.61	3.36
Informer平均值	36.58	3.07	44.89	3.39
LSTM 1倍	85.37	6.31	81.96	5.97
LSTM 2倍	31.23	4.21	38.23	3.64
LSTM 3倍	22.87	3.03	35.57	3.83
LSTM 4倍	34.74	3.84	28.31	2.89
LSTM 5倍	28.37	3.75	33.18	3.16
LSTM平均值	40.52	4.37	43.45	4.43
Transformer 1倍	81.07	4.61	82.15	5.17
Transformer 2倍	39.61	3.77	47.59	2.34
Transformer 3倍	27.53	2.03	38.26	3.47
Transformer 4倍	37.81	3.54	35.81	3.19
Transformer 5倍	33.41	3.43	31.73	3.06
Transformer平均值	39.61	3.07	47.11	4.35

从表3可以看出，Informer各训练集的绝对误差在5倍训练集中为30.37，在5倍测试集中为36.61，说明预测值与实际值差别不大。经5倍交叉验证后训练集的平均误差百分比为3.07%，测试集的平均误差百分比为3.39%。LSTM模型在训练集中，经5倍交叉验证后的平均误差百分比为4.37%；

在测试集中，经5倍交叉验证后的平均误差百分比为4.43%。Transformer模型在训练集中，经5倍交叉验证后训练集的平均误差百分比为3.07%，测试集的平均误差百分比为4.35%。Informer模型在测试集中的5倍交叉验证的平均误差率低于LSTM和Transformer模型，因此Informer神经网络预测模型具有较高的预测精度和实用价值，可以运用于预测华中三省的农产品物流需求。

Informer模型下2021年河南省的预测值为4185.33，实际值为4048.10，误差为3.389%；湖北省的预测值为2503.64，实际值2421.78误差为3.380%；湖南省的预测值为2933.31，实际值为2836.86，误差为3.340%。三个省份预测值与实际值结果较为接近。证明了该模型的有效性后，对华中三省2023年物流需求进行预测，得到2023年河南省预测值4217.13；湖北省预测值2521.47；湖南省预测值2974.65。2023年三省预测值均高于2021年预测值，故应在2021年物流运输配套设施的基础上，保证物流运输效率并加强物流运输能力，从而满足华中地区不断增长的物流需求。

4.3 消融实验

本研究选取了16个影响因素作为自变量，为验证16个指标的选取以及建模参数的科学性，取2021年河南省数据进行消融实验，每次选取15个影响因素作为自变量，将预测值与实际值的绝对误差进行对比，发现16个影响因素作为自变量的绝对误差3.43%低于15个影响因素作为自变量的绝对误差3.51%。如表4所示。

表4 基于河南省2021年数据对16个指标选取是否科学进行验证的消融实验结果

Table 4 Ablation experiment results to verify whether the selection of 16 indicators was scientific based on the data of Henan province in 2021

序号	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	X	Y	绝对误差/%
1	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	3.39
2	×	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	3.52
3	√	×	√	√	√	√	√	√	√	√	√	√	√	√	√	√	3.58
4	√	√	×	√	√	√	√	√	√	√	√	√	√	√	√	√	3.81
5	√	√	√	×	√	√	√	√	√	√	√	√	√	√	√	√	3.77
6	√	√	√	√	×	√	√	√	√	√	√	√	√	√	√	√	3.86
7	√	√	√	√	√	×	√	√	√	√	√	√	√	√	√	√	3.46
8	√	√	√	√	√	√	×	√	√	√	√	√	√	√	√	√	4.07
9	√	√	√	√	√	√	√	×	√	√	√	√	√	√	√	√	3.44
10	√	√	√	√	√	√	√	√	×	√	√	√	√	√	√	√	3.47
11	√	√	√	√	√	√	√	√	√	×	√	√	√	√	√	√	3.62
12	√	√	√	√	√	√	√	√	√	√	×	√	√	√	√	√	3.53
13	√	√	√	√	√	√	√	√	√	√	√	×	√	√	√	√	3.74
14	√	√	√	√	√	√	√	√	√	√	√	√	×	√	√	√	3.48
15	√	√	√	√	√	√	√	√	√	√	√	√	√	×	√	√	3.51
16	√	√	√	√	√	√	√	√	√	√	√	√	√	√	×	√	3.43
17	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	×	3.52

5 结 论

本研究构建Informer神经网络模型作为实验组并选取LSTM和Transformer神经网络作为对照组，对华中地区河南省、湖北省和湖南省农产品物流需求进行预测，选取三省2021年农产品物流需求的实际值和预测值进行对比分析。结果表明，本研究的Informer神经网络模型优于LSTM和Transformer

神经网络模型，可以在农产品物流需求预测中取得较好效果，具有一定的实用价值，可以为华中三省的农产品物流的合理规划提供依据。

通过2023年河南省、湖北省和湖南省的预测结果推测未来的物流需求将会进一步增长，因此华中三省为了满足物流需求需要进一步落实基础设施建设以及完善应急物流协调机制等措施。

本研究存在一定局限性，目前选取的数据指标较多，数据索引难度较高，分析工作量较大，未来应当注重数据的简化性。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 姚冠新, 杨妍, 徐静, 等. 新冠疫情背景下农产品物流需求预测——以兴化市为例[J]. 物流科技, 2021, 44(9): 91-95.
YAO G X, YANG Y, XU J, et al. Forecasting of logistics demand for agricultural products under the background of COVID-19—A case study of Xinghua city[J]. Logistics sci-tech, 2021, 44(9): 91-95.

[2] 张继良. 基于组合模型的山东生鲜农产品物流需求预测[D]. 济南: 山东大学, 2021.
ZHANG J L. Logistics demand forecast of fresh agricultural products in Shandong province based on combination model[D]. Jinan: Shandong University, 2021.

[3] 王晓平, 闫飞. 基于GA-BP模型的北京城镇农产品冷链物流需求预测[J]. 数学的实践与认识, 2019, 49(21): 17-27.
WANG X P, YAN F. Prediction of cold chain logistics demand for agricultural products in Beijing based on GA-BP model[J]. Mathematics in practice and theory, 2019, 49(21): 17-27.

[4] SHYLA M K, SHIVA K K, KUMAR D R. Image steganography using genetic algorithm for cover image selection and embedding[J]. Soft computing letters, 2021, 3: ID 100021.

[5] 皇甫红姣. 绵阳市生鲜农产品冷链物流需求预测研究[D]. 绵阳: 西南科技大学, 2021.
HUANGFU H J. Research on cold chain logistics demand forecast of fresh agricultural products in Mianyang city[D]. Mianyang: Southwest University of Science and Technology, 2021.

[6] 刘文博. 辽宁省生鲜农产品冷链物流需求预测研究[J]. 全国流通经济, 2022(4): 11-13.
LIU W B. Research on demand forecast of cold chain logistics of fresh agricultural products in Liaoning province[J]. China circulation economy, 2022(4): 11-13.

[7] WANG M, LI X. Demand forecasting of agricultural cold chain logistics based on metabolic GM (1, 1) model[J]. IOP conference series: Earth and environmental science, 2021, 831(1): ID 012018.

[8] HUANG L, XIE G, ZHAO W, et al. Regional logistics demand forecasting: A BP neural network approach[J]. Complex & Intelligent Systems, 2021: 1-16.

[9] 徐晓燕, 杨慧敏, 吕修凯, 等. 基于山东省不同模型的物流需求预测比较研究[J]. 包装工程, 2022, 43(23): 207-215.
XU X Y, YANG H M, LYU X K, et al. Comparative research on forecast of logistics demand in Shandong Province based on different models[J]. Packaging engineering, 2022, 43(23): 207-215.

[10] VAN DER LAAN E, VAN DALEN J, ROHRMOSER M, et al. Demand forecasting and order planning for humanitarian logistics: An empirical assessment[J]. Journal of operations management, 2016, 45: 114-122.

[11] LU B, PARK N. A study on model of regional logistics requirements prediction[J]. Journal of Korean navigation and port research, 2012, 36: 553-559.

[12] BAISARIYEV M, BAKYTZHANULY A, SERIK Y, et al. Demand forecasting methods for spare parts logistics for aviation: A real-world implementation of the Bootstrap method[J]. Procedia manufacturing, 2021, 55: 500-506.

[13] ZENG M L, LIU R M, GAO M, et al. Demand forecasting for rural E-commerce logistics: A gray prediction model based on weakening buffer operator[J]. 2022, 2022: ID 3395757.

[14] ANSHUKA A, CHANDRA R, BUZACOTT A J V, et al. Spatio temporal hydrological extreme forecasting framework using LSTM deep learning model[J]. Stochastic environmental research and risk assessment, 2022, 36(10): 3467-3485.

[15] LIU J J, YUAN X, ZENG J H, et al. Ensemble streamflow forecasting over a cascade reservoir catchment with integrated hydrometeorological modeling and machine learning[J]. Hydrology and earth system sciences, 2021, 26(2): 265-278.

[16] XU C W, WANG Y Z, FU H, et al. Comprehensive analysis for long-term hydrological simulation by deep learning techniques and remote sensing[J]. Frontiers in earth science, 2022, 10: ID 875145.

[17] DU S D, LI T R, YANG Y, et al. Deep air quality forecasting using hybrid deep learning framework[J/OL]. arXiv: 1812.04783, 2018.

[18] JANARTHANAN R, PARTHEEBAN P, SOMASUNDARAM K, et al. A deep learning approach for prediction of air quality index in a metropolitan city[J]. Sustainable cities and society, 2021, 67: ID 102720.

[19] PENG J H, CAO H L, ALI Z, et al. Intelligent reflecting surface-assisted interference mitigation with deep reinforcement learning for radio astronomy[J]. IEEE antennas and wireless propagation letters, 2022, 21(9): 1757-1761.

[20] SEDAGHAT N, ROMANIELLO M, CARRICK J E, et al. Machines learn to infer stellar parameters just by looking

at a large number of spectra[J]. Monthly notices of the Royal Astronomical Society, 2021, 501(4): 6026-6041.

[21] LIU H L, LIN P, ZHENG W P, et al. A global eddy-resolving ocean forecast system in China-LICOM Forecast System (LFS)[J]. Journal of operational oceanography, 2021, 16(1): 15-27.

[22] ZHU X M, ZU Z Q, REN S H, et al. Improvements in the regional South China Sea operational oceanography forecasting system (SCSOFSv2) [J]. Geoscientific model development, 2022, 15(3): 995-1015.

[23] 综合百科 [EB/OL]. (2020-10-29) [2023-01-30]. https://www.yebaike.com/22/1574584.html.

[24] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[J/OL]. arXiv:2012.07436 [cs.LG], 2020.

[25] 宋勇, 蔡志平. 大数据环境下基于信息论的入侵检测数据归一化方法[J]. 武汉大学学报(理学版), 2018, 64(2): 121-126.

SONG Y, CAI Z P. Normalized method of intrusion detection data based on information theory in big data environment[J]. Journal of Wuhan university (natural science edition), 2018, 64(2): 121-126.

Forecast and Analysis of Agricultural Products Logistics Demand Based on Informer Neural Network: Take the Central China Aera as An Example

ZUO Min, HU Tianyu, DONG Wei*, ZHANG Kexin, ZHANG Qingchuan

(*College of E-commerce and Logistics*, *Beijing Technology and Business University*, *Beijing* 100048, *China*)

Abstract: Ensuring the stability of agricultural products logistics is the key to ensuring people's livelihood. The forecast of agricultural products logistics demand is an important guarantee for rational planning of agricultural products logistics stability. However, the forecasting of agricultural products logistics demand is actually complicated, and it will be affected by various factors in the forecasting process. Therefore, in order to ensure the accuracy of forecasting the logistics demand of agricultural products, many influencing factors need to be considered. In this study, the logistics demand of agricultural products is taken as the research object, relevant indicators from 2017 to 2021 were selected as characteristic independent variables and a neural network model for forecasting the logistics demand of agricultural products was constructed by using Informer neural network. Taking Henan province, Hubei province and Hunan province in Central China as examples, the logistics demands of agricultural products in the three provinces were predicted. At the same time, long short-term memory network (LSTM) and Transformer neural network were used to forecast the demand of agricultural products logistics in three provinces of Central China, and the prediction results of the three models were compared. The results showed that the average percentage of prediction test error based on Informer neural network model constructed in this study was 3.39%, which was lower than that of LSTM and Transformer neural network models of 4.43% and 4.35%. The predicted value of Informer neural network model for three provinces was close to the actual value. The predicted value of Henan province in 2021 was 4185.33, the actual value was 4048.10, and the error was 3.389%. The predicted value of Hubei province in 2021 was 2503.64, the actual value was 2421.78, and the error was 3.380%. The predicted value of Hunan province in 2021 was 2933.31, the actual value was 2836.86, and the error was 3.340%. Therefore, it showed that the model can accurately predict the demand of agricultural products logistics in three provinces of Central China, and can provide a basis for rational planning and policy making of agricultural products logistics. Finally, the model and parameters were used to predict the logistics demand of agricultural products in Henan, Hunan, and Hubei provinces in 2023, and the predicted value of Henan province in 2023 was 4217.13; Hubei province was 2521.47, and Hunan province was 2974.65, respectively. The predicted values for the three provinces in 2023 are higher than the predicted values in 2021. Therefore, based on the logistics and transportation supporting facilities in 2021, it is necessary to ensure logistics and transportation efficiency and strengthen logistics and transportation capacity, so as to meet the growing logistics demand in Central China.

Key words: agricultural products; demand forecast; Informer neural network; long short-term memory network (LSTM); Transformer; deep learning